COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

LEVEL II

1

TR-996
DAAG-53-76C-0138

January 1981

RELATIVE DEPTH AND LOCAL SURFACE ORIENTATION FROM IMAGE MOTIONS

K. Prazdny

Computer Vision Laboratory
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

A simple mathematical formalism is presented suggesting a mechanism for computing relative depth of any two texture elements characterized by the same relative motion parameters. The method is based on a ratio of a function of the angular velocities of the projecting rays corresponding to the two texture elements. The angular velocity of a ray cannot, however, be computed directly from the instantaneous characterization of motion of a "retinal" point. It is shown how it can be obtained from the (linear) velocity of the image element on the projection surface and the first time derivative of its direction vector. A similar analysis produces a set of equations which directly yield local surface orientation relative to a given visual direction. The variables involved are scalar quantities directly measurable on the projection surface but, unlike the case of relative depth, the direction of (instantaneous) motion has to be computed by different means before the method can be applied. The relative merits of the two formalisms are briefly discussed.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA order 3206) is gratefully acknowledged, as is the help of Janet Salzman in preparing this paper.

1. Introduction

Optical flow is the distribution of angular velocities of projecting rays due to relative motion of objects with respect to the observer. Conceptually, optical flows undoubtedly carry a wealth of information about the spatial arrangement of the viewed scene, and prominent psychologists such as Gibson (1950, 1979) have argued forcefully for the predominant role of this information in human vision. Discontinuities in the distribution of angular velocities have been shown to directly correspond to occluding (or self-occluding) edges (Nakayama and Loomis, 1974). Corresponding discontinuities in "retinal" motions thus offer powerful information for segmentation purposes. Some recent work has illuminated some of the relationships between variables directly involved in the formation of optical flows (Koenderink and van Doorn, 1975, 1976, 1977; Longuet-Higgins and Prazdny, 1980; Prazdny, 1980).

The purpose of this paper is to present a mathematical analysis of some relations containing information about spatial dispositions of a set of texture elements. Using the concept of polar projection as the model for the physical image-forming process we show that the relative depth of two texture elements can be computed as a simple ratio. The entities involved are the angular velocities of the rays through the texture elements and the center of projection, and the visual directions of the rays, which are unit vectors specifying the directions of the rays in

some egocentric reference frame centered at the center of projection. We show that the angular velocity at an image location can be obtained from the image velocity vector and its first time derivative at that locus.

A slightly different analysis requiring an a priori knowledge of the direction vector of the translatory component of the relative motion leads to an interesting characterization of local surface orientation.

Underlying our research on the interpretation of image motions is the assumption that (an approximation to) the velocity vectors associated with individual image elements can be obtained with reasonable accuracy. Some recent research regarding the computation of "retinal" velocities directly from the image brightness values (Hadani et al. 1980; Horn and Schunck, 1980) supports this assumption. Other promising support comes from the research on discrete solutions to the correspondence problem (Ullman, 1979; Barnard and Thompson, 1980) which relies on matching various higher-level image "token" structures.

2. The basic image forming geometry

To begin, let us first consider motions of the projecting rays independently of any particular projection surface. Refer to Figure 1. A texture element P projects, along the ray OP, into the image element Q on the unit sphere centered at the center of projection, O. As the object moves relative to the observer, the ray OP (instantaneously) rotates about an axis through O causing the image element Q to trace a path (T3) on the unit sphere.

Consider a unit vector ($\overline{Q}$) determining the direction of the ray OP at a given instant <Note 1>. Its velocity is $d/dt(\overline{Q})=\dot{\overline{Q}}$. $\overline{Q}$ is perpendicular to $\dot{\overline{Q}}$. As the ray moves with angular velocity $\underline{A}$, Q moves along T3 with (linear) velocity $\dot{\overline{Q}}$ (=$\underline{v}$'). The two velocities are related by $\dot{\overline{Q}}=\underline{A}\times\overline{Q}$. Later we will show how the velocity $\underline{v}$" of an image element on the projection plane relates to the velocity $\dot{\overline{Q}}$ of the unit vector $\overline{Q}$. For the moment it suffices to note that the equation $\dot{\overline{Q}}=\underline{A}\times\overline{Q}$ does not determine $\underline{A}$ uniquely; it only constrains $\underline{A}$ to lie in the plane a normal of which is $\dot{\overline{Q}}$ (see Figure 8 for further explanation).

The motions of the object and/or the observer are relative, i.e., we can only resolve the motion of the object relative to the observer (or vice versa). The (instantaneous) motion of an object with respect to the observer (in the reference frame in which the observer is stationary) can always be described as a rotation (with some angular velocity $\underline{A}_R$) superimposed on a translation specified by a vector $\underline{v}$. The axis of rotation can be

chosen, without loss of generality, to pass through the center of projection (Chasles' theorem) to remove the ambiguity (Whittaker, 1944). The total (linear) velocity of an environmental point P (with position vector $\underline{P}$) is then $\underline{V}=\underline{v}+\underline{A}\times\underline{P}$. An equivalent expression in terms of the angular velocity of the projecting ray OP is

(1) $\underline{A}=\underline{A}_T+\underline{A}_R$

where $\underline{A}_T$ is the angular velocity due to the translation alone, and $\underline{A}_R$ is the rotational component (note that the values of $\underline{A}_R$ and $\underline{v}$ specify relative motion, and not, in any sense, the actual 3D motion of the object). The simplicity of equation (1) results from the fact that angular velocities about a common point add vectorially (Weatherburn, 1965). Observe that $\underline{A}_R$ does not vary from point to point on a rigid body; it is a property of the body as a whole and independent of distance or visual direction. $\underline{A}_T$, on the other hand, is a function of visual direction and the distance of the texture element to the center of projection (see below). Of these two component fields, only $\underline{A}_T$ carries information about relative distance. We will not derive an expression for $\underline{A}_T$ here. The reader is referred, for example, to Nakayama and Loomis (1974) or Prazdny (1980) for a detailed discussion. Briefly, when the object translates relative to the observer, individual rays of projection all move in the same plane (different for different rays). The direction of $\underline{A}_T$ is normal to this plane spanned by the vectors $\bar{v}$ and $\bar{Q}$. The magnitude of $\underline{A}_T$ is equal to $d\beta/dt$ where β is the angle between the direction of translation and the given

ray (see Figure 2). $\underline{A}_T$ is then given by

(2) $\underline{A}_T = d\beta/dt \; (\bar{v}\times\bar{Q})/\sin(\beta) = v/S \; (\bar{v}\times\bar{Q})$

where $S=|OP|$ is the distance of a given texture element to the center of projection.

The observation that enables us to derive an expression for the relative depth of any two texture points moving in the same way relative to the observer concerns equation (1). Consider any two points P_i, P_j on the same object <Note 2>. From equation (1) it follows that

(3) $\underline{A}_i=\underline{A}_{Ti}+\underline{A}_R$ and $\underline{A}_j=\underline{A}_{Tj}+\underline{A}_R$

We see that because $\underline{A}_R$ is the same for the two points it cancels out when the angular velocities are subtracted:

(4) $\underline{A}_{ij}=\underline{A}_i-\underline{A}_j=\underline{A}_{Ti}-\underline{A}_{Tj}$

Using (2) and substituting we obtain

(5) $\underline{A}_{ij}=\bar{v}\times(k_i\bar{Q}_i-k_j\bar{Q}_j)$

where $k_i=v/S_i$. We form the scalar product of both sides of (4) with $(k_i\bar{Q}_i-k_j\bar{Q}_j)$ to obtain

(6) $k_i(\underline{A}_{ij}.\bar{Q}_i)-k_j(\underline{A}_{ij}.\bar{Q}_j)=0$

This is because the scalar triple product involving one vector twice is always zero. We set $a_i=\underline{A}_{ij}.\bar{Q}_i$ and substitute back into (5):

(7) $a_i/a_j=S_i/S_j$

In other words, the relative depth of any two points having the same relative motion parameters ($\underline{v}$ and $\underline{A}_R$) is computable as a simple ratio. Observe that (6) does not involve $\underline{A}_R$ or $\underline{v}$; relative

depth can be computed independently of relative motion. This is a simple but important finding. In general, all mathematical formalisms for computing surface orientation or 3D structure from motions on the projection surface depend (implicitly or explicitly) on computing the rotational component of the relative motion first (e.g., Ullman, 1979; Longuet-Higgins and Prazdny, 1980; Prazdny, 1980). To obtain $\underline{A}_R$, one has to solve a set of nonlinear equations whose coefficients are the velocity vector components of at least five neighboring image elements (Prazdny, 1980), or use the first and second spatial derivatives of the image velocity field to obtain enough information to solve for $\underline{A}_R$ and $\underline{v}$ directly as an integral step in computing the local surface orientation (Longuet-Higgins and Prazdny, 1980). Ullman's scheme (Ullman, 1979), which uses orthogonal projection and relies on a theorem from affine geometry (the structure-from-motion theorem <Note 3>), uses not only spatial information (mutual position of a set of image elements) but also temporal information (the relative position of a given image element in successive snapshots) to recover the rotation of the configuration prior to the computation of relative depth [see Meiri (1980) for some comments on the number of image points and snapshots necessary to solve for the relative motion parameters]. Of course, the assumption of orthogonal projection works only for some situations.

In contrast, equation (6) above only requires that the angular velocities $(\underline{A}_i, \underline{A}_j)$ at two visual directions $(\bar{Q}_i, \bar{Q}_j)$ be known. Unfortunately the angular velocity at a "retinal" locus cannot be computed from the information available at that locus

at an instant. The equation $\underline{v}'=\underline{A}\times\overline{Q}$ does not specify $\underline{A}$ uniquely (see also Figure 8). The angular velocity of a ray through an image point can be obtained only when some additional information is available. We show that it can be computed when the vector specifying the time rate of change in the direction of the "retinal" velocity is available.

To see this consider Figure 3. Due to (relative) motion of a texture element, the ray specified by the direction vector $\overline{Q}$ rotates about O so that Q moves on the surface of the unit sphere with some velocity $\underline{v}'=\dot{\overline{Q}}$. To find the instantaneous plane of rotation of $\overline{Q}$ it is sufficient to apply a few concepts from elementary differential geometry. Observe that $\overline{v}'$, the unit vector in the direction of $\dot{\overline{Q}}$, is the unit tangent to the path at Q. This means that $\dot{\overline{v}}'$ is in the direction of the principal normal at Q. Together, $\overline{v}'$ and $\dot{\overline{v}}'$ span the plane on which lies the circle of curvature at Q. In other words, the plane a normal of which is $\overline{v}'\times\dot{\overline{v}}'$ (this vector lies in the direction of the binormal vector at Q) is the plane of instantaneous rotation of $\overline{Q}$, and $\underline{A}$, the angular velocity vector of Q, must lie in the direction of this vector. Observe that here we bring in temporal information to obtain the angular velocity vector. Other kinds of additional information are possible. In the next section we show, for completeness, how $\overline{v}'$ and $\dot{\overline{v}}'$ relate to "retinal" variables when the projection surface is a plane. Then we analyze a method for computing local surface orientation directly, without computing the relative depth map first.

3. Computing the angular velocity of a projecting ray from "retinal" variables

In this section we assume that the projection surface is a plane at unit distance from O. As the projecting ray rotates about O with some angular velocity $\underline{A}$, Q moves with velocity $\underline{v}'$ along T3, and $\underline{Q}=Q\overline{Q}$, the point at which the ray pierces the projection plane, moves with velocity $\underline{v}''$ along T2 (see Figure 4). Observe also that T3 is a *perspective transformation* of T2 and vice versa. For example, if T3 is a circle, T2 could be any conic section. The exact type of the curve will depend on the mutual orientation of the plane containing T3 (determined by the direction of the angular velocity vector $\underline{A}$) and PP. As mentioned above, the angular velocity vector $\underline{A}$ is normal to the instantaneous plane of rotation of the ray OP, i.e., parallel to the binormal vector. Our first task is thus to determine the direction of the binormal vector associated with the motion of $\overline{Q}$. First, we express $\dot{\overline{Q}}=\underline{v}'$ in terms of the velocity of the image element at a given point, and then we will find the direction of $\underline{A}$ as the vector product of $\overline{v}'$ and $\dot{\overline{v}}'$. We will establish a few *interesting* auxiliary relations before proceeding further.

Consider Figure 5. The figure illustrates the fact that the projection of a segment with magnitude v' along a perpendicular to a given line ℓ on another line ℓ_2 which makes an angle λ with line ℓ is $v''=v'Q/(\sin(\lambda)-v'\cos(\lambda))$. We will refer to this equation as the "radial projection equation." Next we will establish a r[illegible] [illegible]rprising fact about the relation between the velocity of a point and the velocity of its projection. We will first

consider, for simplicity, only planar motions, but the relation holds for space motions too (see below). Consider Figure 6. Suppose that the point 'a' moves along a circular trajectory C so that its distance to O remains unchanged. The (infinitesimal) displacement is $d\varphi$. The displacement of the projection of 'a' on ℓ_1 is $\tan(d\varphi)$. The displacement of the projection of 'a' on ℓ_2 is $Q\tan(d\varphi)/[\sin(\lambda)-\cos(\lambda)\tan(d\varphi)]$ (using the radial projection equation). To compute the relation between the velocities on C and ℓ_2, we divide by dt and take the limit as $dt \to 0$:

$$\lim_{dt\to 0}\left[\frac{Q\tan(d\varphi)}{\sin(\lambda)-\cos(\lambda)\tan(d\varphi)}\ \frac{1}{dt}\right] = \dot{\varphi}\ \frac{Q}{\sin(\lambda)}$$

This is because $\varphi(t)$ is a continuous function of t, and $\lim[\tan(x)/x]=1$ as $x\to 0$. This means that when a point 'a' moves along a path with speed v', its projection on the line ℓ_2 moves with speed-

(8) $v''=Qv'/\sin(\lambda)$

In other words, the velocity of the projection of a point moving along a curve is <u>not</u> the projection of the velocity with which the point moves along that curve <Note 10>. This relation holds also in 3D space <Note 4>. Next we will derive the equation which will enable us to express the angular velocity in terms of the image velocities.

Refer to Figure 4. As $\bar{Q}$ moves along T3, its projection $Q\bar{Q}$ on PP moves along T2 with velocity

(9) $\underline{v}''=d/dt(Q\bar{Q})=\dot{Q}\bar{Q}+Q\dot{\bar{Q}}=\dot{Q}\bar{Q}+Q\underline{v}'$

along T2. Observe that $\underline{v}''$, $\underline{v}'$, and $\bar{Q}$ all lie in the same plane

<Note 4>. This means, however, that the following two vector equation hold simultaneously:

(10) $\bar{v}'\times\bar{Q}=(\bar{v}''\times\bar{Q})/\sin(\lambda)$ and $\bar{v}'.\bar{Q}=0$

We can solve for $\bar{v}'$ from these two simultaneous vector equations <Note 6> to obtain

(11) $\bar{v}'=\bar{Q}\times(\bar{v}''\times\bar{Q})/\sin(\lambda)=\csc(\lambda)\bar{v}''-\cot(\lambda)\bar{Q}$

This is the main equation. Note that $\bar{v}'$ is the unit tangent to T3 at $\bar{Q}$. Its first time derivative, $\dot{\bar{v}}'$, thus lies along the principal normal to T3 at $\bar{Q}$. Their vector product in turn specifies the direction of the sought angular velocity vector $\underline{A}$. Differentiating (11) we obtain

(12) $\dot{\bar{v}}'=\csc(\lambda)\dot{\bar{v}}''-\dot{\lambda}\csc(\lambda)\{\cos(\lambda)\bar{v}''-\bar{Q}\}-\cot(\lambda)\underline{v}'$

Taking the vector product of (12) with $\underline{v}'$ and using the relation $\bar{v}'\times\bar{v}=\cos(\lambda)(\bar{v}'\times\bar{Q})$ leads to

(13) $\bar{v}'\times\dot{\bar{v}}''=\csc(\lambda)(\bar{v}'\times\dot{\bar{v}}'')+\dot{\lambda}(\bar{v}'\times\bar{Q})$

Substituting for $\bar{v}'$ from (11), for v' from (8), and simplifying, we finally obtain

(14) $\bar{v}'\times\dot{\bar{v}}'=Q\csc(\lambda)^2(\bar{v}''-\cos(\lambda)Q)\times\dot{\bar{v}}''+Q\dot{\lambda}\csc(\lambda)(\bar{v}''\times\bar{Q})$

Now, $\dot{\lambda}=d/dt(\lambda)$ can be expressed in terms of $\dot{\bar{v}}''$ and the relation between the visual direction $\bar{Q}$ and the projection plane PP (its unit normal). Using $\bar{v}'.\bar{Q}=\cos(\lambda)$ and differentiating we obtain

(15) $\dot{\lambda}=-\csc(\lambda)(\dot{\bar{v}}''.\bar{Q}+\bar{v}''.\dot{\bar{Q}})=-\csc(\lambda)(\dot{\bar{v}}''.\bar{Q}+\bar{v}''.\bar{v}')$

However, $\bar{v}''.\underline{v}'=(\bar{v}'\sin(\lambda)+Q\cos(\lambda)).\underline{v}'=v'\sin(\lambda)$, because $\bar{v}'.\bar{Q}=0$ by definition. Substituting for v' from (8) yields

(16) $\dot{\lambda}=-\csc(\lambda)(\dot{\bar{v}}''.\bar{Q})-v''\sin(\lambda)/Q$

Equations (14) and (16) indicate that the direction of the angular velocity vector at a visual direction $\bar{Q}$ can be determined completely once the image velocity vector $\underline{v}''$ at the locus on PP corresponding to the visual direction $\bar{Q}$, and the first time derivative of the direction vector of $\underline{v}''$, are available. The only other quantities entering the equation are Q and λ, expressing the metrics of the projective system <Note 6>.

Observe that the direction of the vector $\dot{\bar{v}}''$ is known as soon as $\underline{v}''$ is known. Because T2 is a planar motion, $\dot{\bar{v}}''$ lies in PP, and is perpendicular to $\underline{v}''$. Referring to Figure 7, we see that the direction of $\underline{v}''$, $\bar{v}''$, is specified by $\bar{v}''=\cos(\eta)\bar{x}$ + $\sin(\eta)\bar{y}$ where $\bar{x}$ and $\bar{y}$ are a set of mutually perpendicular unit vectors on PP. $\dot{\bar{v}}''$ is thus given by

(17) $\dot{\bar{v}}''=\dot{\eta}[-\sin(\eta)\bar{x} + \cos(\eta)\bar{y}]$

and the magnitude of $\dot{\bar{v}}''$ is $\dot{\eta}$.

Finally, it remains to specify the magnitude of $\underline{A}$. Observe that A, the magnitude of $\underline{A}$, is a function of the mutual orientation of $\underline{A}$ and the direction vector $\bar{Q}$. This is related to the already mentioned fact that $\underline{v}'$ and $\bar{Q}$ do not specify $\underline{A}$ uniquely (see Figure 8). Using $\underline{v}'=\underline{A}\times\bar{Q}$ we see that $\underline{v}'=v'\bar{v}'=A(\bar{A}\times\bar{Q})=A\sin(\omega)\bar{v}'$, and from this it follows directly that

(18) $A=v'/\sin(\omega)=v''\sin(\lambda)/[Q\sin(\omega)]$

Here, ω is the angle between $\underline{A}$ and $\bar{Q}$, and $\cos(\omega)=\bar{A}.\bar{Q}$.

4. Computing local surface orientation

In this section, we analyze a method of computing local surface orientation relative to a given visual direction. An interesting result will be that the directions of angular velocity vectors are not required explicitly. However, we cannot obtain something for nothing; the analysis requires an a priori knowledge of the (instantaneous) direction of motion (the direction of the translatory component of the relative motion).

First, let us express vectors in a cartesian (rectangular) coordinate frame as a function of two angles α and β. Then, for a given visual direction $\bar{Q}(\alpha,\beta)$, we can compute $\partial\underline{A}/\partial\alpha$ and $\partial\underline{A}/\partial\beta$. Using equation (3), it is easy to see that

$$(19) \qquad \partial\underline{A}/\partial\alpha=\partial\underline{A}_T/\partial\alpha \quad \text{and} \quad \partial\underline{A}/\partial\beta=\partial\underline{A}_T/\partial\beta.$$

In other words, the information contained in the gradient of the angular velocity field in the given visual direction is equivalent to the information contained in the gradient of its translatory component. This is not surprising, for as we saw above, only the translatory component carries information about depth relations between the 3D texture elements.

Let us find expressions for $\partial\underline{A}_T/\partial\alpha$ and $\partial\underline{A}_T/\partial\beta$ and analyze them to see how local surface orientation is specified in these expressions. If α and β are chosen such that the vector corresponding to $\alpha=0$ and $\beta=0$ specifies the direction of the translatory component $\underline{v}$, we see that $\bar{A}_T$ does not change as we move on the plane α=const. Using this fact and differentiating (2) with respect to β yields

(20) $\partial \underline{A}_T/\partial\beta = \partial A_T/\partial\beta\ \bar{A}_T$

Now, $\dot{\beta}=v\sin(\beta)/S$ (see Figure 2), and $\partial A_T/\partial\beta=\partial\dot{\beta}/\partial\beta$. Thus,

(21) $\partial\dot{\beta}/\partial\beta = v\cos(\beta)/S-\dot{\beta}S_\beta/S$ where $S_\beta=\frac{\partial S}{\partial\beta}$

Multiplying both sides by $\tan(\beta)$, and recalling that $\dot{\beta}=v\sin(\beta)/S$, we see that

$\tan(\beta)\partial\dot{\beta}/\partial\beta=\dot{\beta}-\dot{\beta}\tan(\beta)\ S_\beta/S$ so that

(22) $S_\beta/S=-1/\dot{\beta}[\partial\dot{\beta}/\partial\beta]+\cot(\beta)$

To derive a similar expression for S_α/S requires a little more ingenuity. Differentiating equation (2) with respect to α yields

(23) $\partial\underline{A}_T/\partial_\alpha=-\sin(\beta)[v/S][S_\alpha/S]\bar{A}_T+[v/S](\bar{v}\times\bar{Q}_\alpha)$

But we have also

(24) $\partial\underline{A}_T/\partial\alpha=\partial(A_T\bar{A}_T)/\partial\alpha=\partial A_T/\partial\alpha\bar{A}_T+A_T\partial\bar{A}_T/\partial\alpha=\partial\dot{\beta}/\partial\alpha\bar{A}_T+\underline{X}$

where $\underline{X}$ is some vector which does not have to be specified in detail (for our purposes). It can be seen immediately that

(25) $\partial\dot{\beta}/\partial\alpha=-[S_\alpha/S]v\ \sin(\beta)/S$

and from this S_α/S is specified as

(26) $S_\alpha/S=-[1/\dot{\beta}]\partial\dot{\beta}/\partial\alpha$

The quantities S_α/S and S_β/S are depth invariant characterizations of (local) surface orientation relative to a particular visual direction $\bar{Q}(\alpha,\beta)$ <Note 8>. In fact, they are directly related to the gradient of the distance. Because of the dependence of this specification of local surface orientation on a particular visual direction (defining the surfaces of constant α and β), two different surface orientations cannot be directly compared. To

do so, one could transform one characterization into another using a simple rotation matrix. It is important to realize that the expressions characterizing the (local) surface orientation in a pure translatory situation hold also in a general situation of a curvilinear motion. The only prerequisite is (as in the pure translation case) that the direction of the instantaneous motion (i.e., the direction of the translatory component of the curvilinear motion) is known. The knowledge of $\bar{v}$ allows us to define, for each "retinal" locus, a direction along which α=const. By projecting the "retinal" velocity vector into this direction we obtain $\dot{\beta}$, and by differentiating these "retinal" velocities along the directions α=const. and β=const. (see Figure 10) we obtain $\partial\dot{\beta}/\partial\alpha$ and $\partial\dot{\beta}/\partial\beta$.

If this process is to be carried out on the projection plane, the best thing to do is to locate the focus of expansion which then determines the lines α=constant as the lines joining the focus of expansion and the particular "retinal" locus. The localization of the focus of expansion (FOE) is a difficult task. One may try to decompose the image velocity field there into its constituents. The rotational component of the angular veloctiy vector at a given instant can be decomposed into two image velocity fields: one, a hyperbolic field (the component vectors being tangent to hyperbolas through given loci) due to rotation about an axis through the center of projection parallel to the projection plane, and the other, a circular field (the component vectors being tangent to circles with centers at the center of the image coordinate

system) due to rotation about the normal to the projection plane. The remaining translational component would consist of vectors defining the focus of expansion (possibly at infinity) as the unique intersection of all straight lines defined by these vectors and the corresponding "retinal" loci. The whole process is essentially a constrained minimalization problem of a function of three variables: the magnitude of the circular component, the direction of the hyperbolic field (specified by a single angle), and the magnitude of the hyperbolic field (see Figure 11). The constraining condition is that the straight lines of the translational component meet at FOE. Note that such a process, if successful, would essentially recover the translational as well as the rotational component of the angular velocity vector $\underline{A}$ (see equation (1)). We are currently trying to solve this problem using a relaxation scheme <Note 9>.

5. Conclusion

In this paper we have shown that the relative depth of any two texture elements moving in the same way relative to the observer can be computed in a simple way from the angular velocities of the corresponding rays of projection. We have illustrated how the required angular velocity at a point on the planar projection surface can easily be computed from the (linear) velocity of the image element at that locus, and the first time derivative of its direction vector. We have also shown that local surface orientation can be obtained rather straightforwardly once the direction of the translatory component of the relative motion is known. The recovery of this direction from the information contained in the distribution of the "retinal" velocities is a rather complicated task. It is hoped that it may be possible to decompose the instantaneous velocity field on the projection plane into its constituents using a relaxation process. Some work on this problem is currently in progress in our laboratory.

The applicability of the method will depend on the accuracy with which the image velocities can be obtained. It remains to be specified how these errors will propagate through the equations and affect the accuracy of the computed relative depth and local surface orientation.

The computation of relative depth and local surface orientation were presented as two distinct processes. This does not have to be so. Local surface orientation may be obtained from a

relative depth map, for example, by simply fitting a plane to a set of relative depth values in a given (small) neighborhood. I believe that the relative depth map is practically a much more useful construct than local surface orientation. Because the available data are noisy, the computation of local surface orientation relying on quantities obtained in a small neighborhood of a "retinal" point is likely to be affected much more than the relative depth of two widely separated points, where the two angular velocities can be obtained much more precisely, e.g., by averaging over a small neighborhood.

References

1. Clocksin, W. F., 1980, Perception of surface slant and edge labels from optical flow: a computational approach. Perception 9, 253-267.

2. Hadani, I., Ishai, G., Gur, M., 1980, Visual stability and space perception in monocular vision. J. Opt. Soc. Am. 70, 60-65.

3. Horn, B. K. P., Schunck, B. G., 1980, Determining optical flow. MIT AI Memo No. 572, Massachusetts Institute of Technology.

4. Gibson, J. J., 1950, The Perception of the Visual World. Boston: Houghton-Mifflin.

5. Gibson, J. J., 1979, The Ecological Approach to Visual Perception. Boston: Houghton-Mifflin.

6. Koenderink, J. J., van Doorn, A. J., 1975, Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. Optica Acta 22, 772-791.

7. Koenderink, J. J., van Doorn, A. J., 1976, Local structure of movement parallax of the plane. J. Opt. Soc. Am. 66, 717-723.

8. Koenderink, J. J., van Doorn, A. J., 1977, How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow. In Hauske and Butenandt (eds.), Kybernetik, pp. 224-247, Munich: Oldenburg.

9. Longuet-Higgins, C., Prazdny, K., 1980, The interpretation of a moving retinal image. Proc. R. Soc. London B-208, 385-397.

10. Meiri, A. Z., 1980, On monocular perception of 3-D moving objects. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-2, 582-583.

11. Prazdny, K., 1980, Egomotion and relative depth map from optical flow. Biol. Cybernetics 36, 87-102.

12. Weatherburn, C. E., 1965, Elementary Vector Analysis. London: Bell & Sons.

13. Whittaker, E. T., 1944, A Treatise on the Analytical Dynamics of Particles and Rigid Bodies. New York: Dover.

Notes

<Note 1>

The following notation is used throughout the paper. If $\underline{v}$ is a vector then $\bar{v}$ is a unit vector in its direction, and v is the magnitude of $\underline{v}$. The scalar product is denoted by ".", and the vector product by "×". All velocities, position vectors, and the associated quantities are functions of time. This is assumed implicitly throughout the paper. $\dot{\underline{v}}$ and $\dot{v}$ denote the time derivatives, as usual. Angular velocities are vectors perpendicular to the plane of (instantaneous) rotation, with magnitude equal to angular speed (radians/sec).

<Note 2>

The texture elements can be on two different objects as long as the objects move in the same way relatively to the observer (i.e., have the same $\underline{v}$ and $\underline{A}_R$). Thus, for example, in the stationary world where the observer is the only moving agent, the relative depth of all texture elements can be recovered using the present method.

<Note 3>

The structure-from-motion theorem states that the relative depth of four non-coplanar points is recoverable from three non-degenerate *orthographic* projections. The mutual orientation of the projection planes has to be determined before the actual relative depth of the four points can be computed. The recovery of the mutual orientation of the projection planes is an integral part of the schema.

<Note 4>

To see this, note that $\underline{v}''=d/dt(Q\bar{Q})=\dot{Q}\bar{Q}+Q\dot{\bar{Q}}$

Now $\dot{\bar{Q}}=\underline{v}'$ and $\underline{v}'.\bar{Q}=0$. Thus $\underline{v}''\times\bar{Q}=Q(\underline{v}'\times\bar{Q})$, i.e., $\underline{v}'$, $\underline{v}''$, and $\bar{Q}$ all lie in the same plane. Setting $\bar{n}$ to be the unit normal of this plane, we have $\underline{v}'\times\bar{Q}=v'\bar{n}$; but also $\underline{v}''\times\bar{Q}=v''\sin(\lambda)\bar{n}$ (see Figure 4). Thus $\bar{n}=(\underline{v}'\times\bar{Q})/v'=(\underline{v}''\times\bar{Q})/v''\sin(\lambda)$. Substituting for $\underline{v}''\times\bar{Q}$ we obtain

$(\underline{v}'\times\bar{Q})/v'=Q(\underline{v}'\times\bar{Q})/v''\sin(\lambda)$ and so $v'=v''\sin(\lambda)/Q$, as stated in (8).

<Note 5>

A set of vector equations of the form

$$\underline{X} \times \underline{A}=\underline{B} \quad \text{and} \quad \underline{X}.\underline{C}=p$$

where $\underline{A}.\underline{C} \neq 0$, has a general solution $\underline{X}=(p\underline{A}+\underline{C}\times\underline{B})/(\underline{A}.\underline{C})$.

<Note 6>

Given the unit normal $\overline{m}$ defining the projection plane PP, these quantities are computable easily as

$$Q=1/(\overline{Q}.\overline{m}), \text{ and } \lambda \text{ is given by } \cos(\lambda)=(\overline{Q}.\overline{v}").$$

<Note 7>

As can be seen in Figure 9, $\overline{A}_T$ is the same for all $\overline{Q}(\alpha,\beta)$ on the plane α=const.; it is the unit normal of this plane. It follows directly that $\frac{\partial \overline{A}_T}{\partial \beta} = 0$.

<Note 8>

Expressions similar to (22) and (26) were also independently obtained by Clocksin (1980) using a different approach.

<Note 9>

While the problem is conceptually rather simple, there are some difficulties relating to the formulation of the criterion function to be minimized. The difficulty is related to the fact that the projection plane is an augmented Euclidean plane, in terms of projective geometry. For example, the translational vector components on the plane may all meet at an (ideal) point at infinity. It is rather difficult to incorporate this condition into a nicely behaving criterion function.

<Note 10>

This result is intuitively rather surprising. It follows directly, however, from the definition of the angular velocity (see Figure 2).

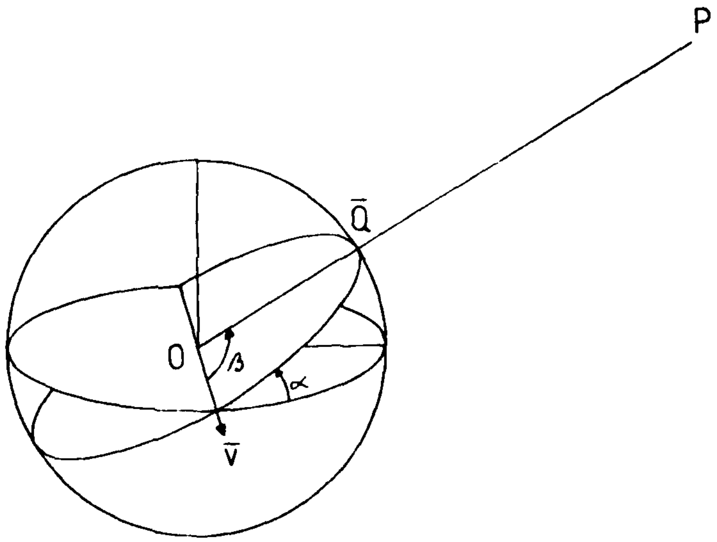

Figure 1. A texture point P projects into a point Q on the unit sphere. The direction vector of the projecting ray OP is determined by the two angles, α, the meridian, and β, the eccentricity; the vector $\bar{Q}$ is a function of α and β. The plane $\alpha=0$ and the direction $(\alpha=0, \beta=0)$ are arbitrary, but it is advantageous for the future analysis to choose them so that the principal x-axis coincides with the direction vector of the translatory motion component.

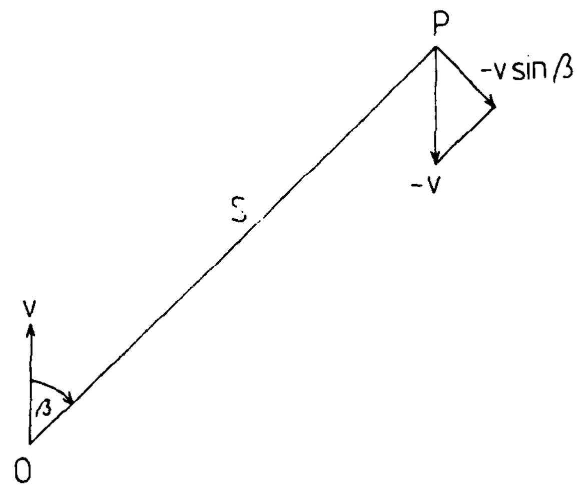

Figure 2. The angular velocity of a ray due to a pure translation. The angular speed is defined as $d\beta/dt$, i.e., as the projection of v on the perpendicular to the ray, divided by the distance S.

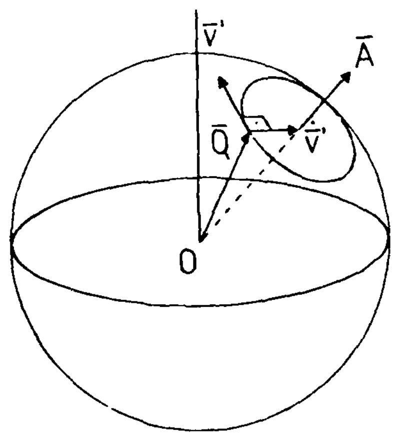

Figure 3. To compute the direction of the angular velocity vector of the ray specified by the visual direction $\bar{Q}$, observe that $\bar{Q}$ moves (on a 3D path on the surface of the unit sphere) with velocity $v'=\bar{Q}$. Because $\bar{v}'$ is the unit tangent to this path, $\bar{v}$ lies in the direction of the principal normal to the path at $\bar{Q}$. The unit binormal vector at $\bar{Q}$, which is perpendicular to the plane spanned by $\bar{v}'$ and $\bar{v}'$, is parallel to the angular velocity vector $\underline{A}$ of $\bar{Q}$.

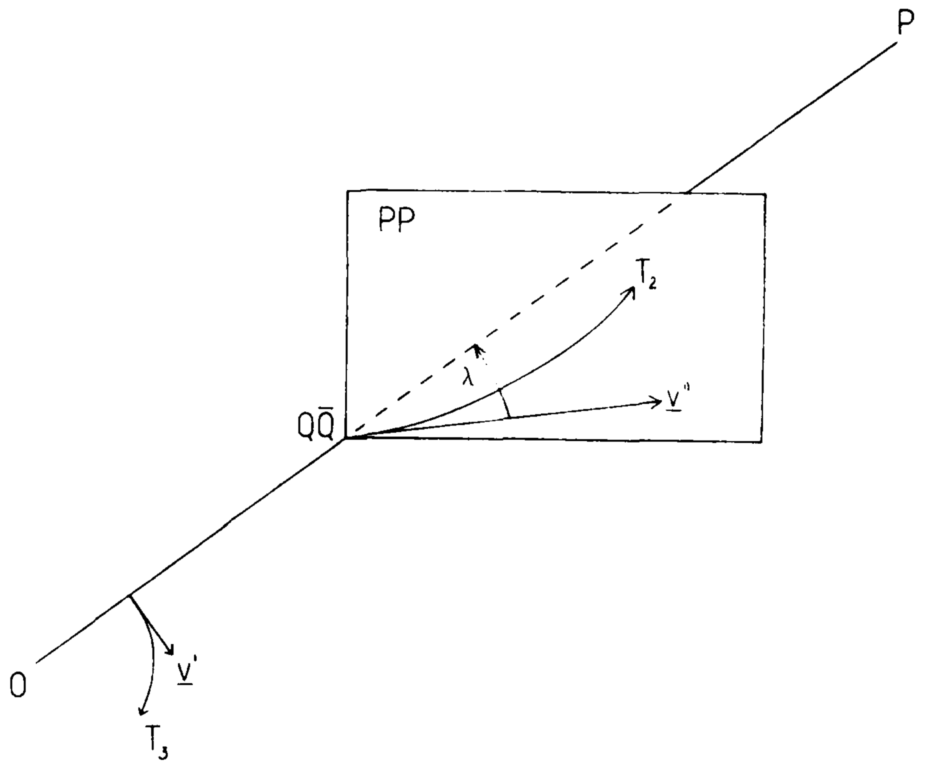

Figure 4. The basic projection geometry. The ray, determined by its direction vector $\bar{Q}$, moves due to the relative motion of the object with respect to the observer. The point $Q\bar{Q}=\underline{Q}$ at which it pierces the planar projection surface, PP, describes a planar trajectory T2. The unit vector $\bar{Q}$ describes a 3D trajectory T3. The angle λ is the angle between the image velocity $\underline{v}''$ of $\underline{Q}$, the projection of P onto PP, and the ray OP.

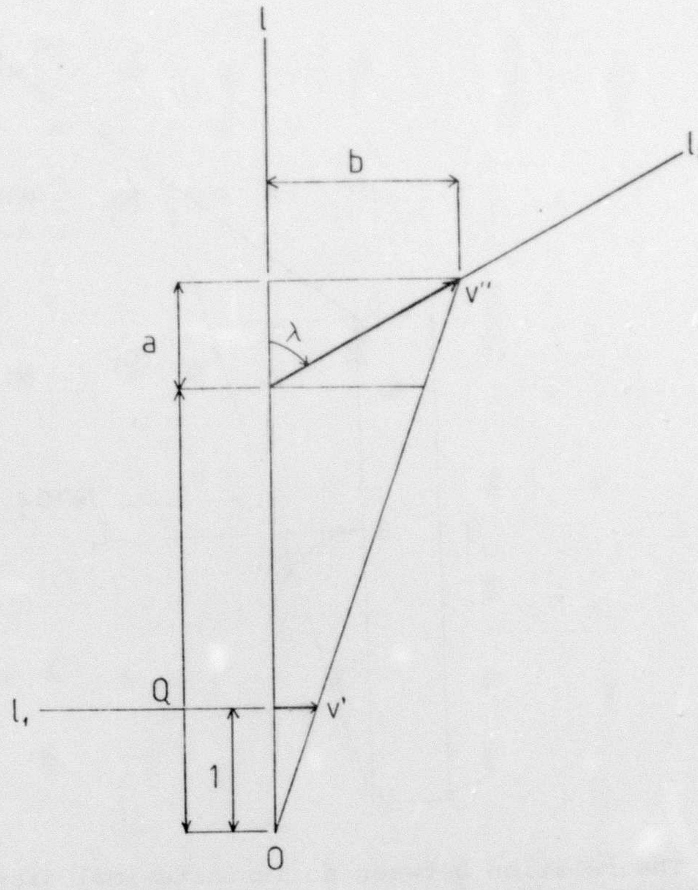

Figure 5. The radial projection equation. The displacement v' on ℓ_1 at unit distance from the center of projection is projected into the displacement v" on ℓ_2. To compute the relation between v' and v", we note that b/(a+Q)=v'; b/v"=sin(λ); and a/v"=cos(λ). Thus v"sin(λ)=v'v"cos(λ)+v'Q. Finally, v"=v'Q/(sin(λ)-v'cos(λ)). Observe that here, v' and v" are finite displacements and not velocity vectors!

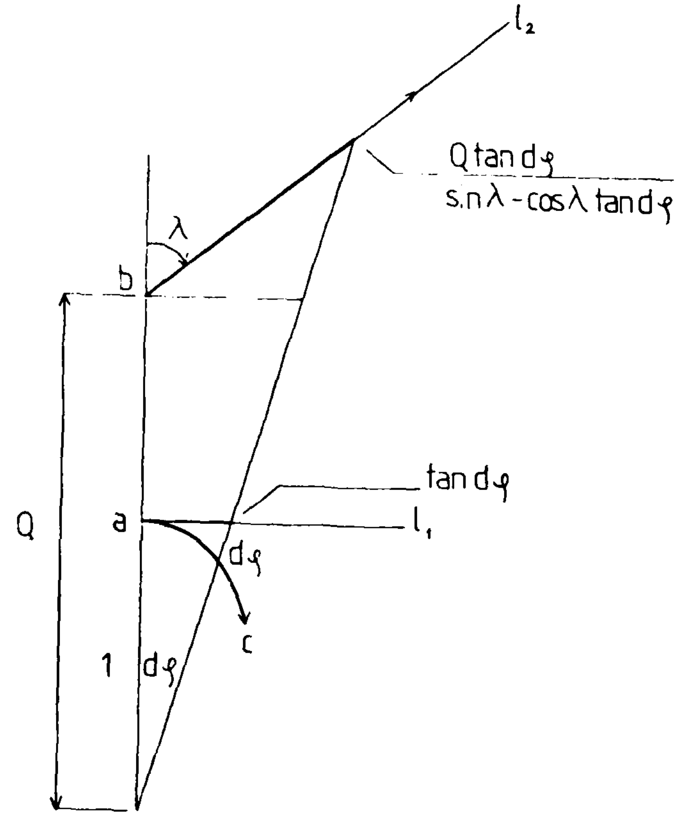

Figure 6. The relation between an infinitesimal displacement $d\varphi$ along the circle C, and its projection on the line ℓ_2. The speed at which b, the projection of a on ℓ_2, moves along ℓ_2 is not the projection of the speed with which a moves along C. See text for further explanation.

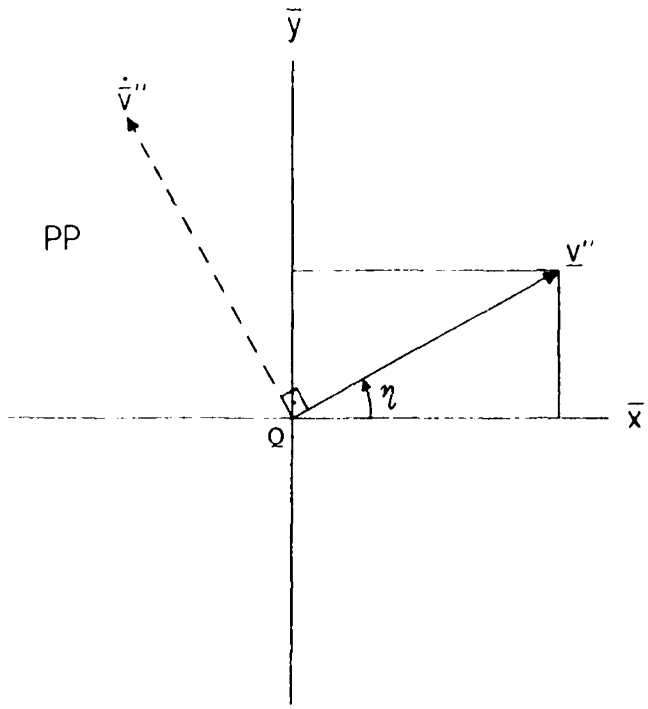

Figure 7. The direction of $\underline{v}''$ on PP is determined by an angle η. The first time derivative of $\underline{v}''$ has direction perpendicular to $\underline{v}''$, and magnitude $\dot{\eta}$. $\bar{x}$ and $\bar{y}$ are a set of mutually perpendicular unit vectors on the projection plane PP.

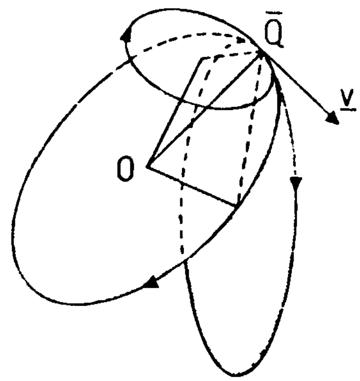

Figure 8. Knowing that the point Q (with position vector $\bar{Q}$) moves with some (linear) velocity $\underline{v}$ does not specify the angular velocity of the ray OQ. The equation $\underline{v}=\underline{A}\times\bar{Q}$ constrains $\underline{A}$ to lie in the plane of which $\underline{v}$ is a normal. Q could (instantaneously) move on an infinite number of possible circles of rotation, only three of them being shown. Observe that A, the magnitude of $\underline{A}$, depends on the angle between $\bar{Q}$ and $\bar{A}$ (it determines the radius of the instantaneous circle of rotation).

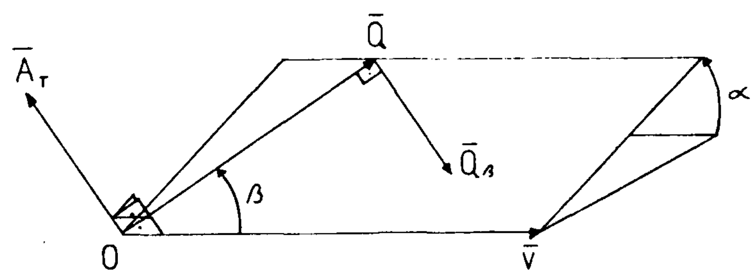

Figure 9. $\bar{A}_T \perp \bar{Q} \wedge \bar{A}_T \perp \bar{Q}_\beta$.

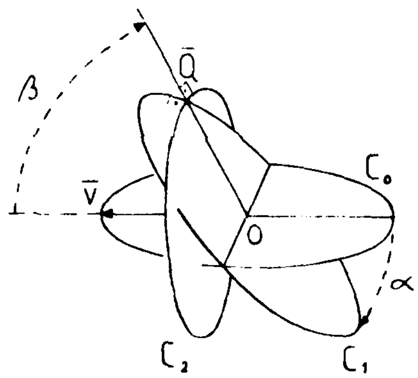

Figure 10. For each given visual direction $\bar{Q}(\alpha,\beta)$, the circles α=const., β=const., and $\bar{Q}$ itself define a rectangular coordinate system. Observe that while the condition α=const. defines a plane, β=const. defines a cone with apex at O! The circles C_1 on the plane and C_2 on the cone are mutually perpendicular.

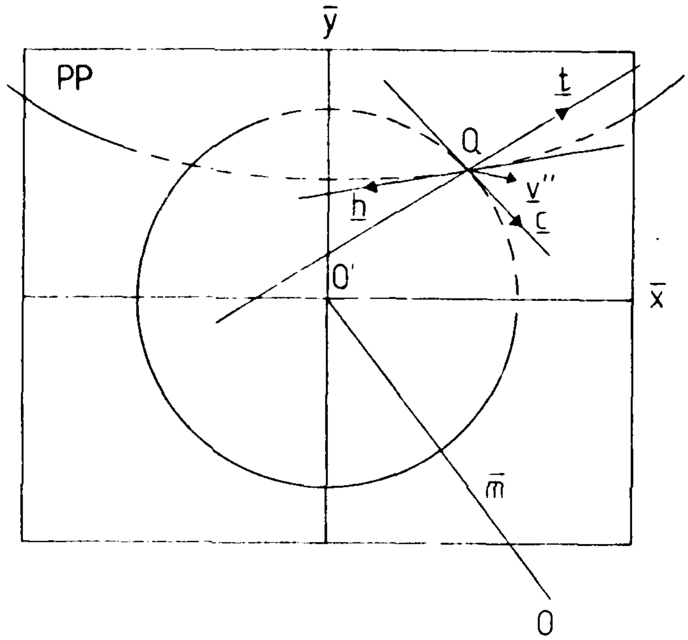

Figure 11. An image velocity $\underline{v}''$ (on the planar projection surface PP) of a point Q can be resolved into three components. The hyperbolic component $\underline{h}$ is due to the rotation of the ray about an axis (through O) in the projection plane (the angular velocity is a linear combination of $\bar{x}$ and $\bar{y}$). The circular component $\underline{c}$ is due to the rotation of the ray about an axis (through O) parallel to $\bar{m}$. The translational component $\underline{t}$ is the remaining vector which constraints the decomposition of $\underline{v}''$; $\underline{t}$ is constrained to be such that $\forall Q_i \in PP$: ($\underline{t}_i$ intersect in one common point). In the illustration above, the direction angle of the hyperbolic field is zero (measured anticlockwise from the x-axis).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. AD-A109 565

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
RELATIVE DEPTH AND LOCAL SURFACE ORIENTATION FROM IMAGE MOTIONS

5. TYPE OF REPORT & PERIOD COVERED
Technical

6. PERFORMING ORG. REPORT NUMBER
(14) TR-996

7. AUTHOR(s)
K. Prazdny (15)

8. CONTRACT OR GRANT NUMBER(s)
DAAG-53-76C-0138, DARPA Order-3206

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Computer Science Center
University of Maryland
College Park, MD 20742

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
U.S. Army Night Vision Lab.
Ft. Belvoir, VA 22060

12. REPORT DATE
January 1981

13. NUMBER OF PAGES
31 (12) 34

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Image processing
Time-varying imagery
Optical flow
Scene analysis
Motion

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
A simple mathematical formalism is presented suggesting a mechanism for computing relative depth of any two texture elements characterized by the same relative motion parameters. The method is based on a ratio of a function of the angular velocities of the projecting rays corresponding to the two texture elements. The angular velocity of a ray cannot, however, be computed directly from the instantaneous characterization of motion of a "retinal" point. It is shown how it can be obtained from the (linear) velocity of the image element on

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 403018

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

the projection surface and the first time derivative of its direction vector. A similar analysis produces a set of equations which directly yield local surface orientation relative to a given visual direction. The variables involved are scalar quantities directly measurable on the projection surface but, unlike the case of relative depth, the direction of (instantaneous) motion has to be computed by different means before the method can be applied. The relative merits of the two formalisms are briefly discussed.